

02D0 #5

OIPE

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/586,106

DATE: 06/14/2000
 TIME: 14:17:51

Input Set : A:\P-10651.app
 Output Set: N:\CRF3\06142000\I586106.raw

ENTERED

3 <110> APPLICANT: Wright, David A.
 4 Voytas, Daniel F.
 6 <120> TITLE OF INVENTION: Plant Retroelements and Methods Related Thereto
 8 <130> FILE REFERENCE: P-1065A
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/586,106
 C--> 11 <141> CURRENT FILING DATE: 2000-06-02
 13 <150> PRIOR APPLICATION NUMBER: 60/087125
 14 <151> PRIOR FILING DATE: 1998-05-29
 16 <150> PRIOR APPLICATION NUMBER: 09/322478
 17 <151> PRIOR FILING DATE: 1999-05-28
 19 <160> NUMBER OF SEQ ID NOS: 165
 21 <170> SOFTWARE: PatentIn Ver. 2.1
 23 <210> SEQ ID NO: 1
 24 <211> LENGTH: 18
 25 <212> TYPE: DNA
 26 <213> ORGANISM: Glycine max
 28 <400> SEQUENCE: 1 18
 29 tggcgccggtt gccaatg
 32 <210> SEQ ID NO: 2
 33 <211> LENGTH: 18
 34 <212> TYPE: DNA
 35 <213> ORGANISM: Glycine max
 37 <400> SEQUENCE: 2 18
 38 tggcgccggtt gtcgggga
 41 <210> SEQ ID NO: 3
 42 <211> LENGTH: 6
 43 <212> TYPE: DNA
 44 <213> ORGANISM: Glycine max
 46 <400> SEQUENCE: 3 6
 47 ttggggg
 50 <210> SEQ ID NO: 4
 51 <211> LENGTH: 7
 52 <212> TYPE: PRT
 53 <213> ORGANISM: Artificial Sequence
 55 <220> FEATURE:
 56 <223> OTHER INFORMATION: Description of Artificial Sequence: plant
 57 retroelement sequence
 59 <400> SEQUENCE: 4
 60 Met Ala Ser Arg Lys Arg Lys
 61 1 5
 64 <210> SEQ ID NO: 5
 65 <211> LENGTH: 1263
 66 <212> TYPE: DNA
 67 <213> ORGANISM: Artificial Sequence
 69 <220> FEATURE:
 70 <223> OTHER INFORMATION: Description of Artificial Sequence: plant
 71 retroelement sequence

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/586,106

DATE: 06/14/2000
 TIME: 14:17:51

Input Set : A:\P-10651.app
 Output Set: N:\CRF3\06142000\I586106.raw

```

73 <400> SEQUENCE: 5
74 atggcctccc gtaaacgcaa agctgtgcc acacccgggg aagcgtccaa ctgggactct 60
75 tcacgtttca ctttcgagat tgcttggcac agataccagg atagcattca gctccggaac 120
76 atccttccag agaggaatgt agagcttgga ccagggatgt ttgatgagtt cctgcaggaa 180
77 ctccagaggg tcagatggga ccaggttctg acccgacttc cagagaagtg gattgatgtt 240
78 gctctggtga aggagtttta ctccaacctt tatgatccag aggaccacag tccgaagtgt 300
79 tggagtgttc gaggacaggt tgtgagattt gatgctgaga cgattaatga ttctctcgac 360
80 accccggtca tcttggcaga gggagaggat tateccagct actctcagta cctcagcact 420
81 cctccagacc atgatcccat cctttccgct ctgtgtactc cagggggacg atttgttctg 480
82 aatgttgata gtgccccctg gaagctgctg cggaaggatc tgatgacgct cgcgcagaca 540
83 tggagtgtgc tctcttattt taaccttgca ctgacttttc acacttctga tattaatgtt 600
84 gacagggccc gactcaatta tggcttgggt atgaagatgg acctggacgt gggcagcctc 660
85 atttctcttc agatcagtc gatcgcccag tccatcactt ccaggcttgg gttcccagcg 720
86 ttgatcaca cactgtgtga gattcagggg gttgtctctg atacctgat ttttgagtca 780
87 ctacgtctcg tgatcaacct tgcctacatt aagaagaact gctggaacct tgccgatcca 840
88 tctatcacat ttcagggggac ccgcgcgacg cgcaccagag ctccgcgctc ggcactctgag 900
89 gctcctcttc catcccagca tccttctcag cctttttccc agagaccacg gcctccactt 960
90 ctatccacct cagcacctcc atacatgcat ggacagatgc tcaggtcctt gtaccagggg 1020
91 cagcagatca tcattcagaa cctgtatcga ttgtccctac atttgcagat ggcactgcca 1080
92 ctcatgactc cggagggccta tcgtcagcag gtcgccaagc taggagacca gccctccact 1140
93 gacagggggg aagagccttc tggagccgct gctactgagg atcctgccgt tgatgaagac 1200
94 ctcatagctg acttggctgg cgctgattgg agcccatggg cagacttggg cagaggcagc 1260
95 tga
98 <210> SEQ ID NO: 6
99 <211> LENGTH: 421
100 <212> TYPE: PRT
101 <213> ORGANISM: Artificial Sequence
103 <220> FEATURE:
104 <223> OTHER INFORMATION: Description of Artificial Sequence: plant
105      retroelement sequence
107 <400> SEQUENCE: 6
108 Met Ala Ser Arg Lys Arg Lys Ala Val Pro Thr Pro Gly Glu Ala Ser
109      1      5      10      15
111 Asn Trp Asp Ser Ser Arg Phe Thr Phe Glu Ile Ala Trp His Arg Tyr
112      20      25      30
114 Gln Asp Ser Ile Gln Leu Arg Asn Ile Leu Pro Glu Arg Asn Val Glu
115      35      40      45
117 Leu Gly Pro Gly Met Phe Asp Glu Phe Leu Gln Glu Leu Gln Arg Leu
118      50      55      60
120 Arg Trp Asp Gln Val Leu Thr Arg Leu Pro Glu Lys Trp Ile Asp Val
121      65      70      75      80
123 Ala Leu Val Lys Glu Phe Tyr Ser Asn Leu Tyr Asp Pro Glu Asp His
124      85      90      95
126 Ser Pro Lys Phe Trp Ser Val Arg Gly Gln Val Val Arg Phe Asp Ala
127      100      105      110
129 Glu Thr Ile Asn Asp Phe Leu Asp Thr Pro Val Ile Leu Ala Glu Gly
130      115      120      125
132 Glu Asp Tyr Pro Ala Tyr Ser Gln Tyr Leu Ser Thr Pro Pro Asp His
133      130      135      140

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/586,106

DATE: 06/14/2000

TIME: 14:17:51

Input Set : A:\P-10651.app

Output Set: N:\CRF3\06142000\I586106.raw

```

135 Asp Ala Ile Leu Ser Ala Leu Cys Thr Pro Gly Gly Arg Phe Val Leu
136 145 150 155
138 Asn Val Asp Ser Ala Pro Trp Lys Leu Leu Arg Lys Asp Leu Met Thr
139 165 170 175
141 Leu Ala Gln Thr Trp Ser Val Leu Ser Tyr Phe Asn Leu Ala Leu Thr
142 180 185 190
144 Phe His Thr Ser Asp Ile Asn Val Asp Arg Ala Arg Leu Asn Tyr Gly
145 195 200 205
147 Leu Val Met Lys Met Asp Leu Asp Val Gly Ser Leu Ile Ser Leu Gln
148 210 215 220
150 Ile Ser Gln Ile Ala Gln Ser Ile Thr Ser Arg Leu Gly Phe Pro Ala
151 225 230 235 240
153 Leu Ile Thr Thr Leu Cys Glu Ile Gln Gly Val Val Ser Asp Thr Leu
154 245 250 255
156 Ile Phe Glu Ser Leu Ser Pro Val Ile Asn Leu Ala Tyr Ile Lys Lys
157 260 265 270
159 Asn Cys Trp Asn Pro Ala Asp Pro Ser Ile Thr Phe Gln Gly Thr Arg
160 275 280 285
162 Arg Thr Arg Thr Arg Ala Ser Ala Ser Ala Ser Glu Ala Pro Leu Pro
163 290 295 300
165 Ser Gln His Pro Ser Gln Pro Phe Ser Gln Arg Pro Arg Pro Pro Leu
166 305 310 315 320
168 Leu Ser Thr Ser Ala Pro Pro Tyr Met His Gly Gln Met Leu Arg Ser
169 325 330 335
171 Leu Tyr Gln Gly Gln Gln Ile Ile Ile Gln Asn Leu Tyr Arg Leu Ser
172 340 345 350
174 Leu His Leu Gln Met Asp Leu Pro Leu Met Thr Pro Glu Ala Tyr Arg
175 355 360 365
177 Gln Gln Val Ala Lys Leu Gly Asp Gln Pro Ser Thr Asp Arg Gly Glu
178 370 375 380
180 Glu Pro Ser Gly Ala Ala Ala Thr Glu Asp Pro Ala Val Asp Glu Asp
181 385 390 395 400
183 Leu Ile Ala Asp Leu Ala Gly Ala Asp Trp Ser Pro Trp Ala Asp Leu
184 405 410 415
186 Gly Arg Gly Ser Glx
187 420
190 <210> SEQ ID NO: 7
191 <211> LENGTH: 1596
192 <212> TYPE: DNA
193 <213> ORGANISM: Artificial Sequence
195 <220> FEATURE:
196 <223> OTHER INFORMATION: Description of Artificial Sequence: plant
197 retroelement sequence
199 <400> SEQUENCE: 7
200 atgcgaggta gaactgcac tggagacgtt gttcctatta acttagaaat tgaagctacg 60
201 tgtcggcgta acaacgctgc aagaagaaga agggagcaag acatagaagg aagtagttac 120
202 acctcacctc ctcttctccc aaattatgct cagatggacg gggaaccggc acaaagagtc 180
203 acactagagg acttctctaa taccaccact cctcagttct ttacaagtat cacaaggccg 240
204 gaagtccaag cagatctcct tactcaaggg aacctcttcc atggtcttcc aatgaagat 300

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/586,106

DATE: 06/14/2000
 TIME: 14:17:51

Input Set : A:\P-10651.app
 Output Set: N:\CRF3\06142000\I586106.raw

```

205 ccatatgcgc atctagcctc atacatagag atatgcagca ccgttaaaat cgccggagtt 360
206 ccaaaagatg cgatactcct taacctcttt tccttttccc tagcaggaga ggcaaaaaga 420
207 tgggtgcact ccttttaaagg caatagctta agaacatggg aagaagtagt ggaaaaattc 480
208 ttaaagaagt atttcccaga gtcaaagacc gtcgaacgaa agatggagat ttcttatttc 540
209 catcaatttc tggatgaatc ccttagcgaa gcactagacc atttccacgg attgctaaga 600
210 aaaacaccaa cacacagata cagcgaagcca gtacaactaa acatattcat cgatgacttg 660
211 caactcttaa tcgaaacagc tactagaggg aagatcaagc tgaagactcc cgaagaagcg 720
212 atggagctcg tcgagaacat ggcggctagc gatcaagcaa tccttcatga tcacacttat 780
213 gttcccacaa aaagaagcct cttggagctt agcacgcagg acgcaacttt ggtacaaaaa 840
214 aagctgttga cgaggcagat agaagccctc atcgaaaccc tcagcaagct gcctcaacaa 900
215 ttacaagcga taagtcttct ccactcttct gttttgcagg tagaagaatg ccccatatgc 960
216 agaggagcac atgagcctgg acaatgtgca agccaacaag acccctctcg tgaagtaaat 1020
217 tatataggca tactaaatcg ttacggattt cagggtctaca accagggaaa tccatctgga 1080
218 ttcaatcaag gggcaacaag atttaatcac gagccaccgg ggtttaatca aggaagaaac 1140
219 ttcatgcaag gctcaagttg gacgaataaa ggaaatcaat ataaggagca aaggaaccaa 1200
220 ccaccatacc agccaccata ccagcaccct agccaaggtc cgaatcagca agaaaagccc 1260
221 accaaaatag aggaactgct gctgcaattc atcaaggaga caagatcaca tcaaaaagagc 1320
222 acggatgcag ccattcggaa tctagaagtt caaatgggcc aactggcgca tgacaaagcc 1380
223 gaacggccca ctagaacttt cgggtgctaac atggagagaa gaaccccaag gaaggataaa 1440
224 gcagtactga ctagagggca gagaagagcg caggaggagg gtaagggttg aggagaagac 1500
225 tggccagaag aaggaaggac agagaagaca gaagaagaag agaaggtggc agaagaacct 1560
226 aagcgtacca agagccagag agcaagggaa gccaaag 1596
229 <210> SEQ ID NO: 8
230 <211> LENGTH: 532
231 <212> TYPE: PRT
232 <213> ORGANISM: Artificial Sequence
234 <220> FEATURE:
235 <223> OTHER INFORMATION: Description of Artificial Sequence: plant
236      retroelement sequence
238 <400> SEQUENCE: 8
239 Met Arg Gly Arg Thr Ala Ser Gly Asp Val Val Pro Ile Asn Leu Glu
240      1          5          10          15
242 Ile Glu Ala Thr Cys Arg Arg Asn Asn Ala Ala Arg Arg Arg Glu
243      20          25          30
245 Gln Asp Ile Glu Gly Ser Ser Tyr Thr Ser Pro Pro Ser Pro Asn
246      35          40          45
248 Tyr Ala Gln Met Asp Gly Glu Pro Ala Gln Arg Val Thr Leu Glu Asp
249      50          55          60
251 Phe Ser Asn Thr Thr Thr Pro Gln Phe Phe Thr Ser Ile Thr Arg Pro
252      65          70          75          80
254 Glu Val Gln Ala Asp Leu Leu Thr Gln Gly Asn Leu Phe His Gly Leu
255      85          90          95
257 Pro Asn Glu Asp Pro Tyr Ala His Leu Ala Ser Tyr Ile Glu Ile Cys
258      100         105         110
260 Ser Thr Val Lys Ile Ala Gly Val Pro Lys Asp Ala Ile Leu Leu Asn
261      115         120         125
263 Leu Phe Ser Phe Ser Leu Ala Gly Glu Ala Lys Arg Trp Leu His Ser
264      130         135         140
266 Phe Lys Gly Asn Ser Leu Arg Thr Trp Glu Glu Val Val Glu Lys Phe

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/586,106

DATE: 06/14/2000

TIME: 14:17:51

Input Set : A:\P-10651.app

Output Set: N:\CRF3\06142000\I586106.raw

```

267 145          150          155          160
269 Leu Lys Lys Tyr Phe Pro Glu Ser Lys Thr Val Glu Arg Lys Met Glu
270          165          170          175
272 Ile Ser Tyr Phe His Gln Phe Leu Asp Glu Ser Leu Ser Glu Ala Leu
273          180          185          190
275 Asp His Phe His Gly Leu Leu Arg Lys Thr Pro Thr His Arg Tyr Ser
276          195          200          205
278 Glu Pro Val Gln Leu Asn Ile Phe Ile Asp Asp Leu Gln Leu Ile
279          210          215          220
281 Glu Thr Ala Thr Arg Gly Lys Ile Lys Leu Lys Thr Pro Glu Glu Ala
282 225          230          235          240
284 Met Glu Leu Val Glu Asn Met Ala Ala Ser Asp Gln Ala Ile Leu His
285          245          250          255
287 Asp His Thr Tyr Val Pro Thr Lys Arg Ser Leu Leu Glu Leu Ser Thr
288          260          265          270
290 Gln Asp Ala Thr Leu Val Gln Asn Lys Leu Leu Thr Arg Gln Ile Glu
291          275          280          285
293 Ala Leu Ile Glu Thr Leu Ser Lys Leu Pro Gln Gln Leu Gln Ala Ile
294          290          295          300
296 Ser Ser Ser His Ser Ser Val Leu Gln Val Glu Glu Cys Pro Thr Cys
297 305          310          315          320
299 Arg Gly Thr His Glu Pro Gly Gln Cys Ala Ser Gln Gln Asp Pro Ser
300          325          330          335
302 Arg Glu Val Asn Tyr Ile Gly Ile Leu Asn Arg Tyr Gly Phe Gln Gly
303          340          345          350
305 Tyr Asn Gln Gly Asn Pro Ser Gly Phe Asn Gln Gly Ala Thr Arg Phe
306          355          360          365
308 Asn His Glu Pro Pro Gly Phe Asn Gln Gly Arg Asn Phe Met Gln Gly
309          370          375          380
311 Ser Ser Trp Thr Asn Lys Gly Asn Gln Tyr Lys Glu Gln Arg Asn Gln
312 385          390          395          400
314 Pro Pro Tyr Gln Pro Pro Tyr Gln His Pro Ser Gln Gly Pro Asn Gln
315          405          410          415
317 Gln Glu Lys Pro Thr Lys Ile Glu Glu Leu Leu Leu Gln Phe Ile Lys
318          420          425          430
320 Glu Thr Arg Ser His Gln Lys Ser Thr Asp Ala Ala Ile Arg Asn Leu
321          435          440          445
323 Glu Val Gln Met Gly Gln Leu Ala His Asp Lys Ala Glu Arg Pro Thr
324          450          455          460
326 Arg Thr Phe Gly Ala Asn Met Glu Arg Arg Thr Pro Arg Lys Asp Lys
327 465          470          475          480
329 Ala Val Leu Thr Arg Gly Gln Arg Arg Ala Gln Glu Glu Gly Lys Val
330          485          490          495
332 Glu Gly Glu Asp Trp Pro Glu Glu Gly Arg Thr Glu Lys Thr Glu Glu
333          500          505          510
335 Glu Glu Lys Val Ala Glu Glu Pro Lys Arg Thr Lys Ser Gln Arg Ala
336          515          520          525
338 Arg Glu Ala Lys
339          530

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/586,106

DATE: 06/14/2000

TIME: 14:17:52

Input Set : A:\P-10651.app

Output Set: N:\CRF3\06142000\I586106.raw

L:10 M:270 C: Current Application Number differs, Replaced Application Number
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date